UNIVERSIDADE TIRADENTES

CURSO DE GRADUAÇÃO DE SISTEMAS DE INFORMAÇÃO

ATIVIDADE PRÁTICA SUPERVISIONADA 1 - APS 1

Estatística

HELENA PASSOS FARIAS

SARA SANTOS DE MELO

LUIZ GUSTAVO MOURA SANTOS

GILVAN DOS SANTOS MOURA JUNIOR

VICTOR CASTELO BRANCO BRASILEIRO PEREIRA

JOÃO GABRIEL BATISTA DOS SANTOS

Aracaju-SE

2024

UNIVERSIDADE TIRADENTES

CURSO DE GRADUAÇÃO DE SISTEMAS DE INFORMAÇÃO

ATIVIDADE PRÁTICA SUPERVISIONADA 1 - APS 1

Estatística

Relatório apresentado à disciplina Estatística, ministrada pelo Prof. Gustavo Rodrigues Borges, no Curso Sistemas de Informação, da Universidade Tiradentes, como atividade prática supervisionada da 1ª Unidade da turma N02, em março de 2024.

Aracaju-SE

2024

SUMÁRIO

1. Identificar as variáveis envolvidas 4
    1.1 Variável Independente (VI) 4
    1.2 Variável Dependente (VD) 4
2. Cálculo da intensidade e direção da correlação linear: 4
3. Ajuste de um modelo matemático 5
    3.1 Seleção do Modelo 5
    3.2 Estimação dos Parâmetros 6
    3.3 Validação do Modelo 6
    3.4 Interpretação dos Resultados 6
    3.5 Previsão e Extrapolação 6
4. Determinação da qualidade do ajuste do modelo 6
    4.1 Coeficiente de determinação (R^2) 7
    4.2 Cálculo dos Resíduos 7
    4.3 Erro Padrão da Estimativa (EPE) 7
    4.4 Teste F e Valor-p 7
5. Utilização do modelo ajustado para fazer previsões 7
    5.1 Seleção das Condições 8
    5.2 Aplicação do Modelo 8
    5.3 Análise das Previsões 8
6. Discussão comparativa 8
    6.1 Comparação dos Coeficientes 8
    6.2 Análise dos (R^2) 9
    6.3 Exame dos Resíduos 9
    6.4 Erro Padrão da Estimativa 9
    6.5 Validade Estatística 9
    6.6 Discussão dos Resultados 9
7. Principais conclusões 9
REFERÊNCIAS 10

- **1. Identificar as variáveis envolvidas**

  Em um experimento ou estudo observacional, as variáveis são classificadas em dois tipos principais:

  1.1 Variável Independente (VI): É a variável que é manipulada ou controlada pelo pesquisador para testar seus efeitos sobre a variável dependente. Em outras palavras, é a causa presumida que influencia ou determina o valor da variável dependente. Por exemplo, se você está testando o efeito de diferentes quantidades de água no crescimento das plantas, a quantidade de água é a variável independente.

  1.2 Variável Dependente (VD): É a variável que é medida ou observada para ver como ela é afetada pela variável independente. É o efeito ou resultado que está sendo estudado. No exemplo anterior, o crescimento das plantas seria a variável dependente, pois é o que você mede para ver como ele muda em resposta às diferentes quantidades de água aplicadas.

  Para identificar essas variáveis em um processo ou experimento, deve-se primeiro entender qual é o objetivo do estudo e quais são os fatores que estão sendo testados ou observados. A variável independente é aquela que você altera intencionalmente para ver se há uma mudança na variável dependente, enquanto a variável dependente é aquela cuja mudança você está interessado em medir como resultado das alterações na variável independente.

- **2. Cálculo da intensidade e direção da correlação linear:**

  Para calcular a intensidade e a direção da correlação linear entre duas variáveis, usamos o coeficiente de correlação de Pearson, também conhecido simplesmente como (r). O coeficiente de correlação de Pearson mede o grau de relação linear entre duas variáveis quantitativas. A fórmula para calcular (r) é:

  Fórmula:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

  Exemplo: Suponha que após calcular o coeficiente de correlação de Pearson, você obtenha um valor de (r = 0.65), indicando uma correlação

positiva moderada entre o tamanho do conteúdo da página e o tempo de carregamento.

Onde:

(x_i) e (y_i) são os valores individuais das variáveis,

(\bar{x}) e (\bar{y}) são as médias dos valores das variáveis **(x) e (y)**, respectivamente.

O valor de (r) varia entre -1 e 1:

- Um valor de **(r = 1) indica uma correlação positiva perfeita**, onde os valores de uma variável aumentam exatamente com o aumento dos valores da outra.
- Um valor de **(r = -1) indica uma correlação negativa perfeita**, onde os valores de uma variável diminuem com o aumento dos valores da outra.
- Um valor de **(r = 0) indica que não há correlação linear** entre as variáveis.
- Valores próximos a 0 indicam uma correlação fraca, enquanto valores próximos a -1 ou 1 indicam uma correlação forte.

Ao comentar os resultados, você deve considerar tanto a intensidade quanto a direção da correlação:

- Uma correlação **positiva** significa que, à medida que uma variável aumenta, a outra também tende a aumentar.
- Uma correlação **negativa** significa que, à medida que uma variável aumenta, a outra tende a diminuir.

OBS: É importante lembrar que correlação não implica causalidade. Mesmo que duas variáveis estejam fortemente correlacionadas, isso não significa que uma causa a outra.

- **3. Ajuste de um modelo matemático:**

Ajustar um modelo matemático para descrever o comportamento dos dados experimentais é um processo fundamental em muitas áreas da ciência e engenharia. O objetivo é encontrar uma função matemática que melhor represente a relação entre as variáveis estudadas. Aqui estão os passos básicos para ajustar um modelo matemático:

3.1 Seleção do Modelo: Escolher um modelo matemático que possa descrever bem os dados. Isso pode ser baseado em conhecimento teórico ou observações prévias. Por exemplo, modelos lineares são

comuns, mas às vezes modelos polinomiais, logísticos, exponenciais ou outros podem ser mais adequados.

3.2 Estimação dos Parâmetros: Usar métodos estatísticos para estimar os parâmetros do modelo escolhido. Isso geralmente envolve o uso de técnicas como mínimos quadrados ou máxima verossimilhança para ajustar a linha ou curva aos dados de forma que a diferença entre os valores observados e os valores previstos pelo modelo seja minimizada.

3.3 Validação do Modelo: Após estimar os parâmetros, é importante validar o modelo para garantir que ele descreve adequadamente os dados. Isso pode envolver o cálculo de medidas de qualidade de ajuste (como (R^2)), análise de resíduos, e realização de testes de hipóteses para verificar se os parâmetros são significativamente diferentes de zero.

3.4 Interpretação dos Resultados: Interpretar os parâmetros estimados em termos do contexto do estudo e discutir se o modelo faz sentido com base no conhecimento subjacente do processo ou fenômeno estudado.

3.5 Previsão e Extrapolação: Se o modelo for validado, ele pode ser usado para fazer previsões sobre novos dados ou para extrapolar além do intervalo dos dados experimentais.

- Um modelo linear simples pode ser ajustado para descrever como a largura de banda da rede influencia o tempo de transferência de dados.

- Para ajustar um modelo de regressão linear simples, deve-se procurar uma linha que melhor descreva a relação entre a Variável Dependente (VD) e a Variável Independente (VI).

Fórmula da Regressão:

$$y = mx + b$$

Exemplo: Utilizando o método dos mínimos quadrados, você pode encontrar a equação da linha de regressão linear (y = 0.5x + 1.2), onde (y) é o tempo de carregamento e (x) é o tamanho do conteúdo.

- **4. Determinação da qualidade do ajuste do modelo:**

Para determinar a qualidade do ajuste de um modelo matemático aos dados experimentais, utilizamos várias métricas estatísticas. Segue a explicação de cada uma delas e como podem ser interpretadas:

4.1 Coeficiente de determinação (R^2): Pode ser usado para avaliar o quão bem o modelo se ajusta aos dados experimentais.O valor do coeficiente indica a proporção da variância na variável dependente que é explicada pelo modelo. Um (R^2) de 1 significa que o modelo explica toda a variância, enquanto um (R^2) de 0 indica que o modelo não explica nenhuma variância. Valores mais altos de (R^2) indicam um melhor ajuste do modelo aos dados.

4.2 Cálculo dos Resíduos: Os resíduos são as diferenças entre os valores observados e os valores previstos pelo modelo. Analisar os resíduos pode ajudar a identificar padrões que o modelo não capturou e possíveis outliers.

4.3 Erro Padrão da Estimativa (EPE): Esta métrica fornece uma medida da precisão das previsões do modelo. Um EPE menor indica que as previsões do modelo estão mais próximas dos valores reais.

4.4 Teste F e Valor-p: O teste F verifica se há uma relação significativa entre a variável dependente e a variável independente. Um valor-p baixo (geralmente menor que 0,05) indica que o modelo é estatisticamente significativo.

Ao comentar os resultados, deve-se considerar todas essas métricas em conjunto. Por exemplo, um (R^2) alto junto com um EPE baixo e um valor-p baixo no teste F sugeriria que o modelo tem um bom ajuste. No entanto, é importante também olhar para os resíduos para garantir que não há padrões não explicados ou problemas com o modelo.

Fórmula:

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

Exemplo: Calcule o coeficiente de determinação ((R^2)), que indica a proporção da variação da VD explicada pela VI.

Um (R^2) de 0.72 sugere que 72% da variação do tempo de resposta é explicada pela carga de trabalho.

- **5. Utilização do modelo ajustado para fazer previsões:**

  O modelo ajustado pode ser usado para prever o tempo de resposta para diferentes níveis de carga de trabalho.

Exemplo: Para uma carga de trabalho de 100, 150 e 200 unidades, calcule os tempos de resposta esperados.

Para utilizar um modelo matemático ajustado e fazer previsões para condições não estudadas experimentalmente, deverão ser seguidos os seguintes passos:

5.1 Seleção das Condições: Escolha pelo menos três valores diferentes para a variável independente (VI) que estejam dentro da faixa de dados usada para ajustar o modelo. Esses valores devem ser novos e não devem ter sido usados na análise original.

5.2 Aplicação do Modelo: Use o modelo ajustado e os parâmetros estimados (por exemplo, (m) e (b) na regressão linear) para calcular as previsões da variável dependente (VD) para cada uma das novas condições selecionadas.

5.3 Análise das Previsões: Avalie as previsões geradas pelo modelo. Verifique se elas são consistentes com o que seria esperado com base no conhecimento teórico ou em dados semelhantes.

Por exemplo, se o modelo de regressão linear ajustado é:

$$y = 2x + 3$$

e desejamos fazer previsões para os valores de (x) iguais a 5, 10 e 15, o cálculo seria:

- Para (x = 5): (y = 2(5) + 3 = 13)
- Para (x = 10): (y = 2(10) + 3 = 23)
- Para (x = 15): (y = 2(15) + 3 = 33)

Essas previsões fornecem uma estimativa de (y) para cada nova condição de (x) dentro da faixa de ajuste do modelo.

- **6. Discussão comparativa:**

Para realizar uma discussão comparativa entre a qualidade dos ajustes de um modelo linear feitos pelo autor de um trabalho selecionado na literatura e pelo aluno, deve-se considerar os seguintes pontos:

6.1 Comparação dos Coeficientes: Devemos verificar se os coeficientes do modelo (inclinação e intercepto) são semelhantes entre o trabalho do autor e o do aluno. Diferenças significativas podem indicar variações na metodologia ou nos dados utilizados.

6.2 Análise dos (R^2): Comparar os valores do coeficiente de determinação ((R^2)) de ambos os modelos. Um (R^2) mais alto indica um melhor ajuste do modelo aos dados.

6.3 Exame dos Resíduos: Analisar os padrões dos resíduos (diferenças entre os valores observados e os previstos) para ambos os ajustes. Resíduos aleatórios sugerem um bom ajuste, enquanto padrões nos resíduos podem indicar um modelo inadequado.

6.4 Erro Padrão da Estimativa: Comparar o erro padrão da estimativa de ambos os modelos. Um erro padrão menor indica que as previsões do modelo estão mais próximas dos valores reais.

6.5 Validade Estatística: Avaliar a validade estatística dos modelos comparando os testes F e os valores-p associados aos coeficientes. Modelos com valores-p baixos indicam que os coeficientes são estatisticamente significativos.

6.6 Discussão dos Resultados: Comentar sobre as semelhanças e diferenças encontradas, e tentar explicar as possíveis razões para essas diferenças. Considere aspectos como a seleção de dados, métodos de coleta e análise, e a interpretação dos resultados.

Ao fazer essa comparação, é importante ter acesso aos detalhes completos dos ajustes realizados tanto pelo autor quanto pelo aluno, incluindo os dados utilizados, as fórmulas aplicadas, e as análises estatísticas. Isso permitirá uma avaliação precisa e uma discussão fundamentada sobre a qualidade dos ajustes de modelo linear.

- **7. Principais conclusões:**

  Neste trabalho, propomos a inclusão dos conteúdos de Correlação Linear e o cálculo dos coeficientes da reta de Regressão Linear Simples no conjunto de conteúdos de Matemática do Ensino Médio, uma vez que tais assuntos são uma aplicação direta dos conceitos da Estatística Descritiva. Para tanto, realizamos, num primeiro momento, uma fundamentação teórica a respeito dos principais resultados relacionados à Estatística Descritiva. Em seguida, desenvolvemos os conteúdos de Correlação Linear e Regressão Linear Simples a partir da fundamentação teórica realizada anteriormente. Por fim, apresentamos uma proposta didática aliada a resolução de problemas para o ensino da Correlação e Regressão Lineares utilizando o Software Matemático GeoGebra, que é uma das Tecnologias de Informação e Comunicação (TIC) mais utilizadas nos últimos anos para o Ensino da Matemática.

REFERÊNCIAS

https://ninocarvalho.com/blog/o-que-e-stakeholders-significado-identificacao/

https://edisciplinas.usp.br/pluginfile.php/5625867/mod_folder/content/0/124-449-2-PB.pdf

https://www.academia.edu/89376481/PRESSMAN_Engenharia_de_software_Uma_Abordagem_Profissional_9a_Ed

https://inventti.com.br/5-livros-essenciais-para-profissionais-de-software/

https://embarcados.com.br/correlacao-covariancia-e-determinacao-regressao-linear-parte-2/

https://www.bing.com/search?q=estudos+%C3%A1rea+TI+correla%C3%A7%C3%A3o+linear&toWww=1&redig=E9C4BD0C12F54551960D85B16AD4988D

https://edisciplinas.usp.br/pluginfile.php/4574088/mod_resource/content/1/aula_correlacaoregressao.pdf

https://linkconcursos.com.br/coeficiente-de-correlacao-linear-e-o-significado-dos-resultados/

https://education.ti.com/pt/atividades/portugal/actividades/associacao-linear

https://home.uevora.pt/~cpires/diagnost/regressao2.pdf

https://est.ufmg.br/ftp/fcruz/ep/Cap11.pdf

https://statorials.org/pt/correlacao/

https://cesad.ufs.br/ORBI/public/uploadCatalago/09222427022012Bioestatistica_Aula_10.pdf

https://www.scielo.br/j/acr/a/xxyrq5kGswF8ZfrwrF6Q7nP/

https://www.repositorio.ufop.br/jspui/bitstream/123456789/12961/1/ARTIGO_Aplica%c3%a7%c3%a3oAn%c3%a1liseCorrela%c3%a7%c3%a3o.pdf

https://bing.com/search?q=estudo+com+correla%c3%a7%c3%a3o+linear+simples+em+artigos+acad%c3%aamicos

https://doi.org/10.1590/2317-6431-2019-2143

http://dx.doi.org/10.1016/S0006-3223